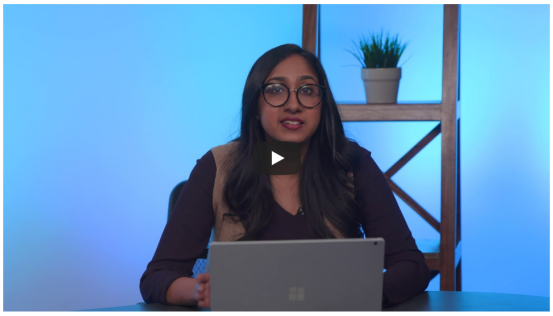
Common Types of Data



It's All Numerical in the End

Note that although we've described *numerical data* as a distinct category, it is actually involved in some way with all of the data types we've described. With the example of stock performance (above) the stock prices are numerical data points. So why do we give this as an example of "time-series data" rather than "numerical data"? It is the ordering of the numerical data points *across points in time* that leads us to call the data *time-series data*.

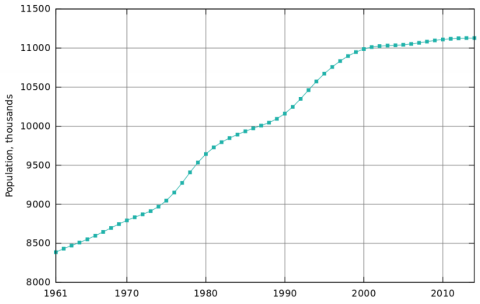
What is more, all data in machine learning eventually ends up being numerical, regardless of whether it is numerical in its original form, so it can be processed by machine learning algorithms.

For example, we may want to use gender information in the dataset to predict if an individual has heart disease. Before we can use this information with a machine learning algorithm, we need to transfer male vs. female into numbers, for instance, 1 means a person is male and 2 means a person is female, so it can be processed. Note here that the value 1 or 2 does not carry any meaning.

Another example would be using pictures uploaded by customers to identify if they are satisfied with the service. Pictures are not initially in numerical form but they will need to be transformed into RGB values, a set of numerical values ranging from 0 to 255, to be processed.

QUESTION 1 OF 2

Have a look at this graph:



Population of Greece since 1961 (*Wikimedia Commons*)

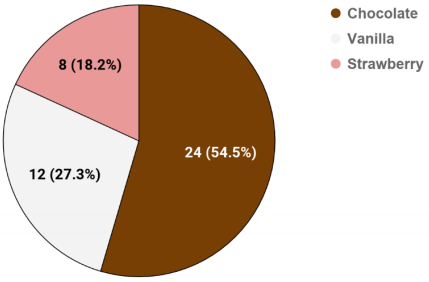
What type of data is this?

- Numerical
- ✓ Time-Series
- Categorical
- Text

SUBMIT

QUESTION 2 OF 2

Have a look at this chart showing the number of people who like each flavor of ice cream:



What type of data is this?

- Numerical
- Time-Series
- ✓ Categorical
- Text

SUBMIT

NEXT